Wax

ARPA Network Information Center
Stanford Research Institute
Menlo Park, California 94025
NIC # 16743
Packet Radio Temporary Note # 52

Network Analysis Corp.
1-June-73

PACKET RADIO BROADCAST NETWORK-SYSTEM OPERATION

I. INTRODUCTION

The operation of an ALOHA type broadcast packet radio network is described in this note. Particular topics addressed are: packet routing and control, acknowledgments, and labelling of repeaters. The general network structure and operation emerged from previous Packet Radio Notes (e.g., [1] and[2] which demonstrated the necessity of having efficient routing schemes, and [3] which shows the advantages of hop-by-hop acknowledgment) and Packet Radio Meetings. The operation described here will be tested by simulation.

The broadcast network interfaces with other networks as shown schematically in Figure 1. The entire system will be composed of a terrestrial point to point communication network of the ARPA type, the terrestrial broadcast network described here, and possibly a network of satellite ground stations which utilize satellite channels in a point-to-point or broadcast mode. A message which originates from a mobile terminal will first arrive to a station and then (if necessary) it can use either the terrestrial network or the satellite network.

The objective of the broadcast network is to accumulate traffic from many mobile terminals which have a high ratio of peak to average bandwidth requirements. Due to the difficulty of traffic routing and control [1], and the practical difficulty of using a high channel bandwidth (> Mb/s), it is desirable that packets arrive to the nearest station and use the shortest possible path.

In this note we concentrate on a system which is composed of one station, a set of repeaters and many terminals. The terminals are assumed to originate messages for transmission to the station and the station responds with a message (several packets) to the originating terminal. A single channel is used for transmission from terminals to station and from station to terminals. Devices access the channel at random points in time, i.e., non-slotted ALOHA mode. Spread spectrum modulation will be used to improve the capture effect and overcome possible interference due to multipath reflection.

A more general system with many stations, and terminal to terminal communication, will operate, by and large, in the same manner. However, additional specifications and requirements will be necessary, some of which are discussed.

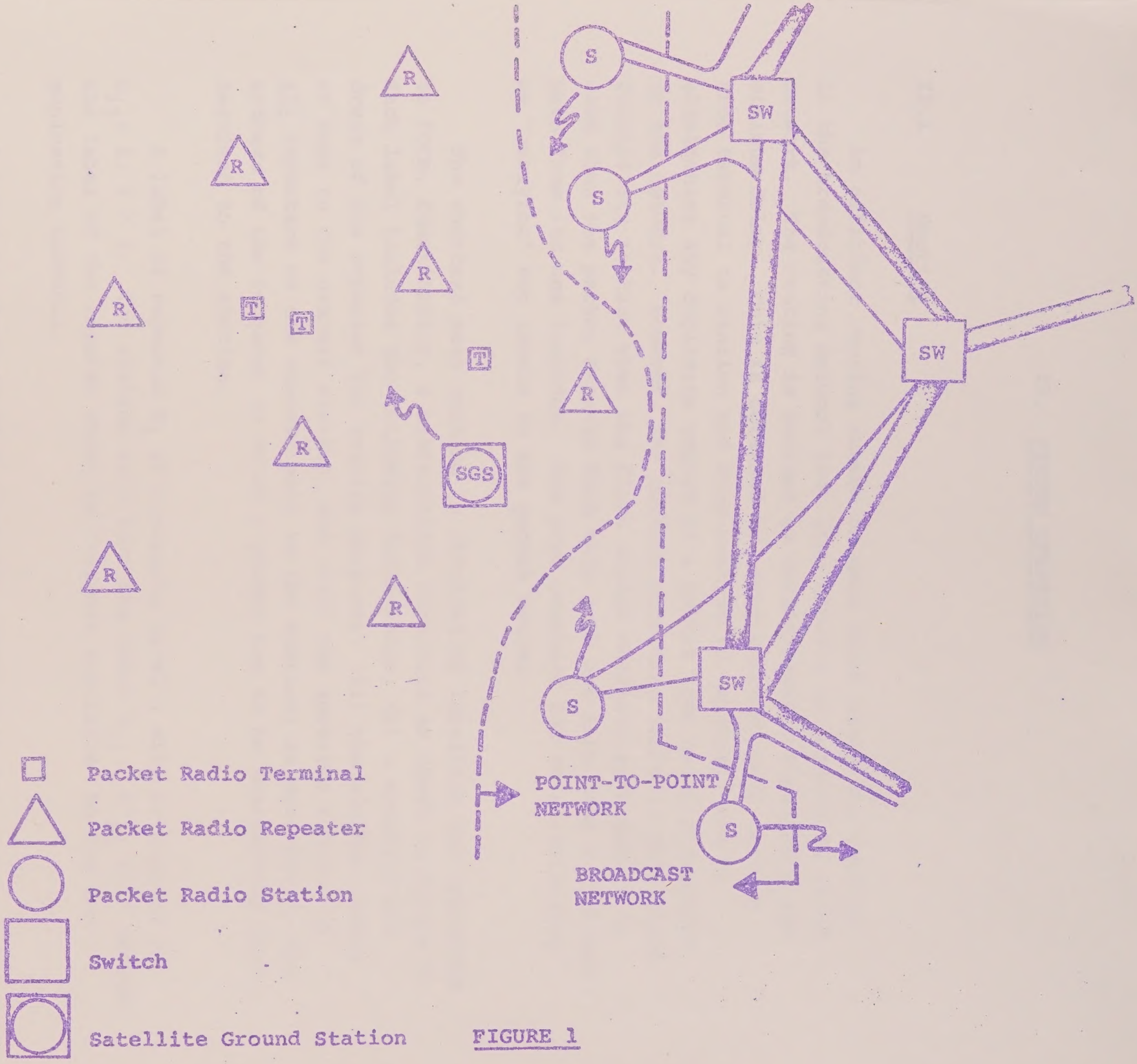

R
R
R
R
R
R
R
T
T
T
SGS
S
S
S
SW
SW
SW
POINT-TO-POINT
NETWORK
BROADCAST
NETWORK
Packet Radio Terminal
Packet Radio Repeater
Packet Radio Station
Switch
Satellite Ground Station

FIGURE 1

II. SYSTEM OPERATION

II.1 Routing

An efficient routing scheme is used which makes the routing in the broadcasting network look line routing in a point to point network. The routing is devised to obtain maximum throughput and minimum delay. This is obtained by using shortest path routing from terminal to station and from station to terminal, and by eliminating any duplicate copies of a packet from being circulated in the network. However, the routing procedure includes sufficient flexibility so that when the first choise shortest path cannot be used then the packet departs form this path and then uses a shortest path from its new location. One pays overhead for this efficiency by "carrying" two labels in the packet header.

The shortest path routing is obtained by labelling the repeaters to form, functionally, a hierarchical structure as shown in Figure 2. Each label includes the following information: (i) a specific address of the repeater for routing purposes, (ii) the minimum number of hops to the nearest station, and (iii) the specific address of all repeaters on the nearest path to the station; in particular, the address of the repeater to which a packet has to be transmitted when destined to the station.

A label of repeater R_i of hierarchy level j will be denoted by L_{ij}; i, j > 1. The station will have the label L_{11}. L°_{ij} will denote the label of the repater which is "nearest available" to the communicating terminal.

A label is composed of H subfields, where H is the maximum number of hierarchy levels (H-1 is the maximum number of hops on the shortest path between any repeater and the station). Every subfield has three possible entries, blank (BLK), a serial number (SER), or ALL. L_{ij} has j entries SER's and (H-j) BLK's as shown below:

Digitized by the Internet Archive
in 2025 with funding from
Amateur Radio Digital Communications, Grant 151

https://archive.org/details/temporary-note-052

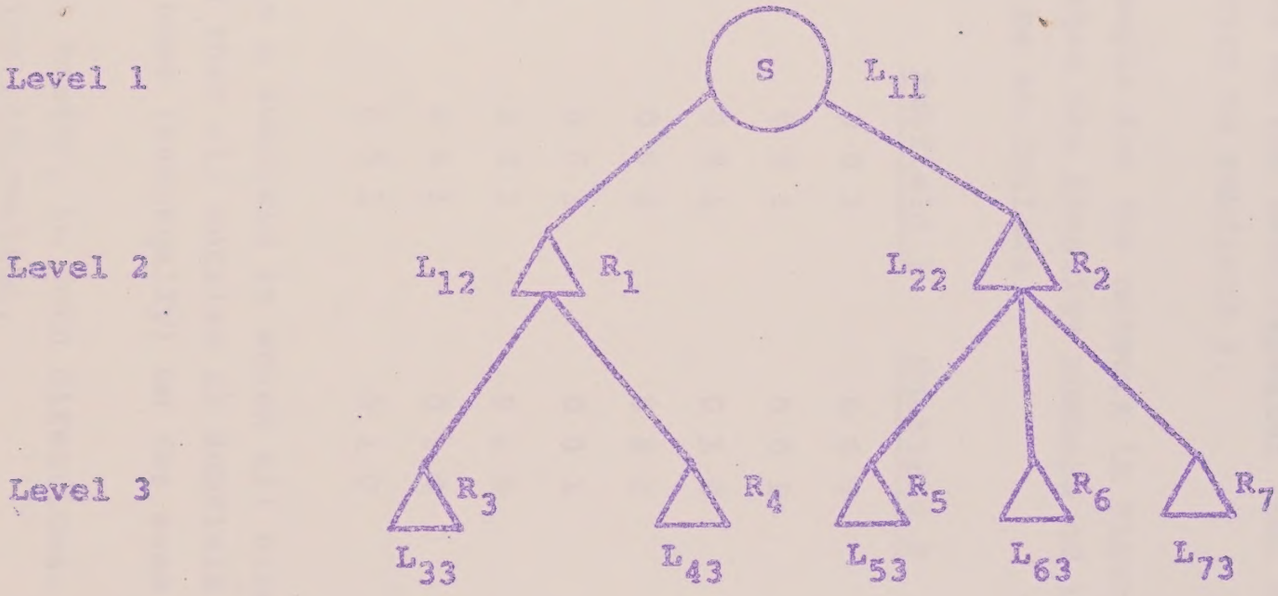
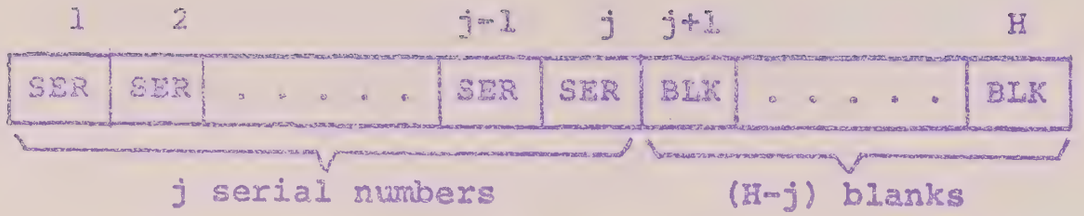

Level 1
Level 2
Level 3
S
L_{11}
L_{12}
R_1
L_{22}
R_2
R_3
R_4
R_5
R_6
R_7
L_{33}
L_{43}
L_{53}
L_{63}
L_{73}

FIGURE 2



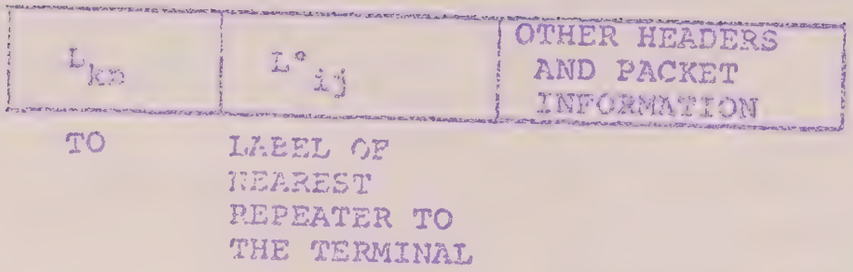
We say that L_{ij} homes on L_{kp}, $h(L_{ij}) = L_{kp}$, if $p = j-1$ and the first j-1 subfields of both are identical. If two repeaters at level j home on the same repeater then their labels will differ only in the entry to subfield j.

As an example for the network in Figure 2; if we take 3 bits per subfield then the identity numbers of the station and the repeaters can be as follows:

	Subfield 1	Subfield 2	Subfield 3
L_{11}	0 0 1	0 0 0	0 0 0
L_{12}	0 0 1	0 0 1	0 0 0
L_{22}	0 0 1	0 1 0	0 0 0
L_{33}	0 0 1	0 0 1	0 0 1
L_{43}	0 0 1	0 0 1	0 1 0
L_{53}	0 0 1	0 1 0	0 0 1
L_{63}	0 0 1	0 1 0	0 1 0
L_{73}	0 0 1	0 1 0	0 1 1

In this example a subfield in which all bits are "0" is considered "blank". Note that all entries in Subfield 1 are the same since all repeaters home (eventually) on the same station.

The packet header, in both directions, will include the following routing information.



L_{kn} is the label of the repeater to which the packet is currently addressed. The complete packet will always be transmitted to a specific device; other devices which may receive the packet will drop it. The shortest path froma terminal to the station consists of L_{ij}, $h(L_{ij})$, $h(h(L_{ij}))$, up to L_{11}, in the given order, and in the reverse order from station to terminal. When a specific repeater along the shortest path is not known (by the terminal) or not available, then the terminal or repeater (which have the packet) will transmit only the header part of the packet, trying to identify a specific repeater. In that case, the label L_{kn} will include some entries ALL.

Routing from Terminal to Station: When a mobile terminal begins to communicate, it first identifies a repeater in its area. It transmits only the header part of the packet with all entries in L_{kn} set to ALL. The header is addressed to all repeaters and statin that can hear the terminal. A device which correctly receives this header substitutes its label in the space L_{kn} and repeats the header. This particular L_{kn} is also L°_{kn} and will be used by the terminal to transmit all packets during this period of communication. If a terminal is not mobile, then it can save this label for future transmissions. L°_{kn} begins to transmit the complete packet along the shortest path to the station.

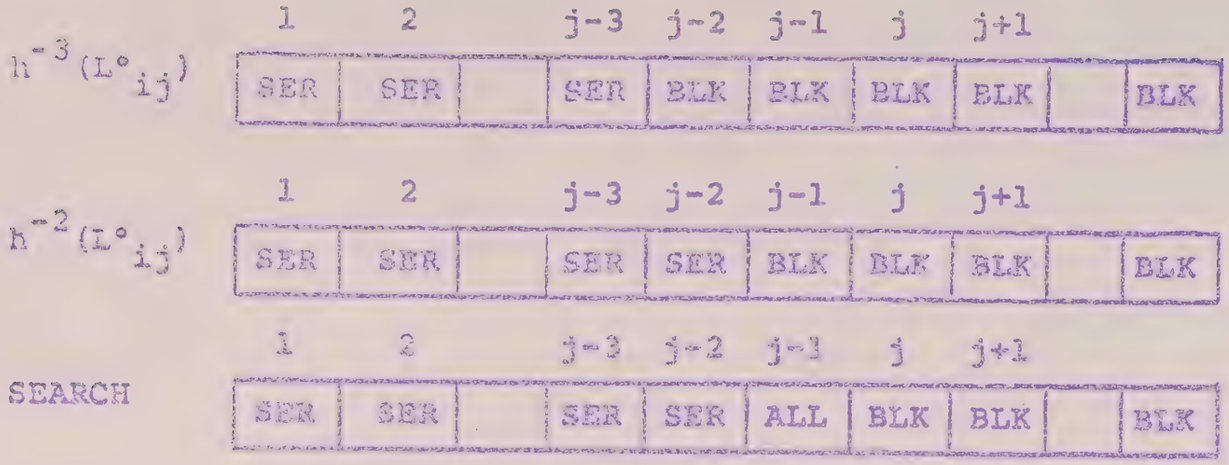
Suppose that L_{ij} along the shortest path is not successful in transmitting the packet to $h(L_{ij})$. Then L_{ij} begins the search stage of trying to identify another repeater. In the first step it tries to identify a repeater which is in level $p \leq j-1$. This is done by using the following label:

1	2	3		j-1	j	j+1		
SER	ALL	ALL	. . .	ALL	BLK	BLK	. . .	BLK

The header is addressed to all repeaters in levels 2 to j-1 which eventually home on L_{11}. If this step is not successful then in the second (last) step L_{ij} tries to identify any available repeater by using a label in which the first entry is SER and all other entries are ALL. When a specific repeater is identified and receives the packet, it transmits the packet on the shortest path from its new location.

Note that when repeaters have processing capability and storage they can save alternative labels and thus reduce the necessity of searching for a specific repeater. Alternative solutions in which repeaters have multiple labels are also possible.

Routing from Station to Terminal: L°_{ij} contains sufficient information for shortest path routing to the terminal. Denote by h^{-1} the inverse of h and by $h^{-2} = h^{-1}(h^{-1})$, etc. The shortest path from station to terminal includes $h^{-(j-1)}(L^{\circ}_{ij})$, $h^{-(j-2)}(L^{\circ}_{ij})$, ... $h^{-1}(L^{\circ}_{ij})$, and L°_{ij}. If some L_{kp} is not successful in transmitting the packet along the shortest path then it begins the process of identifying another specific repeater. Note that when routing to the station then the next label is always a function of the label of the repeater that currently stores the packet; whereas when routing to the terminal then the next label is a function of L_{ij} and the hierarchy level of the repeater that currently stores the packet. Thus when routing to the terminals, it will be useless to transmit the packet backwards since it will usually arrive back to the current location, so it will be more efficient to delay the packet. Therefore, when routing to the terminal and some repeater on the shortest path is temporarily out of order, then the procedure tries to by-pass this particular repeater and then get back onto the the original shortest path. We show the labels that will be used by $h^{-3}(L^{\circ}_{ij})$.



All the entries SER are taken from L°_{ij}.

II.2 Acknowledgments

The acknowledgment scheme will be Hop-byHop Positive Echo Acknowledgment. That is, there will be a positive acknowledgment on the last hop (in each direction) where the packet arrives to its destination and not repeated. Along the path there will be echo acknowledgment, namely, the device which transmits the packet waits in a receive mode to receive this same packet when it is repeated by the next stage. The reception of the packet when transmitted by the next stage constitutes the acknowledgment since it indicates that the next repeater stage has correctly received the packet and will store and retransmit it as necessary. In fact, to identify the above it is sufficient to "hear" the header of the packet, thus, the header plus header parity bits will constitute the acknowledgment. At the end of a path, the terminal or station will repeat the header. Note that the probability of correctly receiving an acknowledgment is higher than the probability of correctly receiving a packet due to the difference in transmission time. Furthermore, in many cases the station may correctly receive th header, whereas the entire packet is received in error. The information contained in the header can be used by the station for control purposes.

An end-to-end positive acknowledgment may also be needed. However, if this is used after each packet than it reduces the throughput, but more severely it increases the average delay per message [3]. Initially, we will use end-to-end acknowledgment at the end of the message. Packets will be numbered and assimilated at the station (or possibly at the destination station). The first packet will be marked "initial" and have the count of the number of packets in the message; if one of the packets in the sequence is not received at the station, then the station will send an error indication to the terminal requesting the retransmission from that number on [4]. Thus, it constitutes a kind of an end-to-end negative acknowledgment plus positive acknowledgment at the end of the message.

Remark:

In a multi-hop system, one cannot use hop-by-hop negative acknowledgment, since in many cases when a packet has not been correctly received that the receiving device may not know where the packet was transmitted from (when omni-directional antennas are used).

II.3 Transmission and Retransmission

All devices will use non-slotted ALOHA random access mode. The first transmission as well as the retransmissions will be randomized over time. Terminals and repeaters will keep retransmitting the same packet until acknowledged or until the maximum number of retransmissions specified is reached. The station may alternate in packets transmitted since they may be directed to different repeaters.

When a device completes the transmission of a packet, it sets a random time, say t_r, and changes into a receiving mode. t_r will be uniformly distributed between T and, say kT*, where T is the transmission time of a packet. If the device is not echo acknow-

* The value of k is not yet decided and depends whether some of the carrier-sense features will be implemented.

ledged within t_r then it retransmits the packet. The device may change the label in the header when retransmitting, as described in the section on routing. If the device is acknowledged and has another packet for transmission then it can start transmitting the other packet at $t_i > T$ so that no interference with the first transmission of its own packet is possible. When a terminal successfully transmits the last packet of its message, it will remain in a receiving mode, waiting for the end-to-end message acknowledgment or for a request for retransmission of part of the message.

II.4 Traffic Control

To suppress a packet from being indefinitely repeated when one departs from the shortest path and to limit the number of packet transmissions when the channel traffic is high, we use a maximum handover number (MHN) and a Maximum Number of Retransmissions (MNR), respectively.

It was demonstrated in previous PRT Notes that when the channel traffic is beyond a certain value, then the throughput reduces. If the number of retransmissions is not limited, then the channel traffic will increase indefinitely and the throughput will reduce to zero. However, if the channel traffic is reduced to a "stable" level, then after a period of "settling time" the system will return to normal operation. Thus, the MNR is used to enable the system to return to normal operation. If the terminal stops retransmission and turns off after MNR, then the packets will be considered "blocked". However, it may be sufficient that the terminal stop retransmission for a longer period. This can be obtained by having the terminal wait a random time, uniformly distributed over a large interval, say 20T; in this case the packet is delayed.

The MHN will not be constant, but rather depend on the number of hops on the shortest path for the particular packet. It will be assigned by the repeater whose label is L^o_{ij} when a

packet is transmitted to the station, and by the station when the packet is transmitted to a terminal.

One may wish to have higher priority for some control packets or end-to-end acknowledgment packets. Apart from the initial labelling procedure, most of the control packets will be shorter which naturally gives them higher priority due to a lower probability of conflict and shorter transmission time. If this is not sufficient then the following can be introduced: When a repeater identifies a control packet of a specific type, it will store and delay an ordinary information packet (if it has one) and keep on transmitting the control packet.

II.5 Packet Format

The information packet format is the following:

HEADER	PARITY	PACKET INFORMATION	PARITY

The header will include the following items:

T/F	C/I	DID	OID	L_{Km}	L°_{ij}	MHN	E/C	P

T/F - a bit, indicating whether the packet is addressed To station or From station.

C/I - a bit, indicating whether the packet is a Control packet or an Information packet.

DID - Destination address.

OID - Origination address.

L_{Km} - The label of the repeater to which the packet is currently addressed.

L°_{ij} - The label of the repeater "nearest" to the terminal which originated the packet or to which the packet is transmitted.

MHN - Maximum handover number.

E/C - An error or control message. If it is an information packet then the space may include a sequential number, specification of the number of packet in the message, etc. If it is a "header packet" then it may include an error message asking for retransmission of a certain number of packets. If it is a control packet from the station then this space may be used for the control message; etc.

P - Parity bits for the header.

III. A PROCEDURE FOR REPEATER LABELLING

Initially, every station and repeater has a fixed ID, R_i. R_i will be used for labelling purposes, to identify whether the device is operative or dead, to turn a device on or off, and other control purposes.

The station will determine and assign labels to all repeaters in its area in the initial labelling procedure. When more than one station operate in an area, the initial labelling will be done by the stations sequentially, and repeaters may be allowed to choose the home station accoring to the minimum number of hops.

First, it is necessary to specify two parameters: (i) The maximum number of subfields or hierarchy levels, say H, and (ii) the number of bits per subfield, say B. These parameters are to be the same for the entire broadcast network in order to have the same packet format when transmitting information. If some sections of the broadcast network are disjoint then it is sufficient that B x H be the same for the entire network. As indicated before, H-1 is the maximum number of hops that a packet will travel when using the shortest path route, and 2^B-2 is the maximum number of repeaters that can home on a single repeater or station (one label is needed for ALL and another for BLK). 2 x H x B is the number of bits in the header which will contain the routing information.

The initial labelling procedure is done as follows:

Step I:

The station transmits a control packet to every repeater sequentially. This packet includes a MHN as well as another MHN to be used by the addressed repeater for its response packet. There is no routing at this stage and every repeater which correctly receives the control packet decrements its MHN, stores and retransmits it until echo acknowledged by the next stage. The control packet is dropped when its MHN reduces

to zero. The repeater to which this packet is addressed transmits a response packet to the station using the assigned MHN. Every repeater which receives this packet will decrement the MHN and add its R_i in order.

The station may receive one response packet, several, or none. If no response packet is received, then the station can try several more transmissions, each time increasing the MHN's, or decide that the repeater is dead (this repeater can possibly be reached from another station).

Step II:

The information acquired from the response packets is sufficient for determining a hierarchical labelling structure. In this step, the station processes the information and determines a structure which is optimal in some sense. The processing done in this step is described in the next section.

Step III:

In this step the station tests the shortest path, particularly in the direction from station to repeaters which was not tested before. The station transmits a control packet to every repeater, using its label. The station uses a MHN equal to the number of hops on the shortest path so that if this path is not possible then the repeater will not be able to receive the packet. A repeater which receives this packet transmits a response to the station which constitutes a positive acknowledgment. If all repeaters have been successfully tested then the procedure ends; otherwise the program returns to Step II for further processing.

III.2 An Algorithm for Determining the Labels

We describe a technique for processing the response packets in Step II of the labelling procedure and determining the hierarchical labels. In general, the repeaters may be distributed at random locations in some area and the station may not know the

geographical locations of repeaters. Furthermore, there may be more repeaters than the number needed for efficient routing. Ideally, one would want to obtain a network of repeaters which has the following properties:

(1) Minimum number of hierarchy levels.

(2) Shortest path from every repeater to the station.

(3) The entire area should be covered with a minimum number of repeaters.

(4) Every repeater should be able to transmit directly to at least j (say 2) other repeaters.

(5) The number of repeaters which home on one single repeater or station should be $\leq 2^B-2$.

A solution which satisfies all the requirements may not exist. For example, if more than 2^B-2 repeaters can directly reach the station and none of them can be taken out because of area coverage or for satisfying (4), then (2) will not be satisfied.

Suppose that there are N-1 repeaters and one station denoted by R_1. The station first constracts a connectivity matrix $C = (c_{ij})$, where,

$$c_{ij} = \begin{cases} 1 & \text{if device j can hear i directly} \\ 0 & \text{otherwise.} \end{cases}$$

C is constracted from the response packets in Step I. For example, if the response from R_i contains (in order) R_e, R_m, R_k,..., then $c_{ie} = 1$, $c_{em} = 1$, $c_{mk} = 1$,.... One can see that the station does not have to transmit the first labelling packet to all repeaters since it can learn about the functional location of some of the repeaters which were on the return path of previous control packets. Furthermore, the number of response packets to a control packet can be increased when the MNN assigned by the station is increased. Finally, we note that C is not necessarily symmetric.

The entries 1 in row i indicate the repeaters to which R_i can directly transmit, and the entries in column i indicate the repeaters from which R_i can directly receive. The structure of the repeater network will be recorded by the vector h_j, $j = 1,\ldots,N$ where h_j indicates the repeater on which R_j homes.

Let S(m) denote the set of repeaters whose shortest path to the station includes exactly m hops. Assume that all repeaters in S(1), S(2),..., and S(m), have been labelled (assigned home repeaters). We describe the labelling of (repeaters in) S(m+1) by (repeaters of) S(m). At every state k of labelling S(m+1) by S(m) we characterize repeaters of S(m) by:

$d_i(k)$ = The number in S(m+1) which have been labelled R_i (say, the degree of R_i at state k)

$v_i(k)$ = The number in S(m+1) which still can be labelled R_i

$f_i(k)$ = The potential degree of R_i at state k, i.e. $f_i(k) = d_i(k) + v_i(k)$.

Repeaters of S(m+1) will be characterized by:

v_j = The number of repeaters in S(m) which can label it.

At every state k, we distinguish among three disjoint subsets of S(m) and S(m+1):

$S_m^F(k) = \{R_i : v_i(k) = 0\}$ = cannot label more*

$S_m^L(k) = \{R_i : f_i(k) \leq 2^B-2;$ ordered according to increasing values of $d_i(k)$ and when the same then according to increasing values of $v_i(k)\}$

---

* There may be repeaters in S(m) for which $v_i(0) = 0$. It is convenient to refer to these as "end repeaters of level m" since no repeater will home on these.

$S^R_m(k) = \{R_i : f_i(k) > 2^B - 2$; ordered according to increasing values of $v_i(k)\}$

$S^F_{m+1}(k) = \{R_i : h_i(k) > 0\}$ = already labelled (assume that $h_i(0) = 0$)

$S^L_{m+1}(k) = \{R_i : c_{ip} = 1$ for some $R_p \in S^R_m(k)$; ordered according to decreasing values of the number of repeaters in $S^R_m(k)$ that can label them and when this is the same then according to decreasing values of $u_i\}$

$S^R_{m+1}(k)$ = The remaining repeaters of S(m+1) ordered according to decreasing values of u_i

Note that $S^R_m(k)$ is the set which can potentially violate requirement (5) and $S^L_{m+1}(k)$ is the set to be labelled which may result in this violation. Therefore, one should try to label $S^L_{m+1}(k)$ by $S^L_m(k)$; when such a label is assigned, it decreases the values of $f_i(k)$ in $S^R_m(k)$. Furthermore, the orders of the subsets of S(m) according to $d_i(k)$ is aimed to obtain a network in which the repeaters of S(m+1) are divided equally among repeaters of S(m) (this was not specified in the requirements). The order of $S^L_{m+1}(k)$ is done so that first (if possible) is labelled the repeater which can be labelled by the largest number of repeaters of $S^R_m(k)$.

The algorithm proceeds as follows:

Step A:

Take the first of $S^L_{m+1}(k)$ and label it by one of $S^L_m(k)$ (beginning with the first). If it is labelled, evaluate the subsets of state $k+1$ and do the same; if not, take the next repeater of $S^L_{m+1}(k)$ and do the same.

Step B:

(i) If $S^R_m(k)$ is empty then $S^L_{m+1}(k)$ is also empty; complete the labelling of $S^R_{m+1}(k)$ by $S^L_m(k)$ using the procedure of Step A.

(ii) If $S^R_m(k)$ is not empty, label one of $S^L_{m+1}(k)$ by $S^R_m(k)$ using the procedure of Step A, then return to Step A.

Note that if all of $S_m^R(k)$ becomes part of $S_m^L(k)$ at some state then the network produced satisfies requirements (1), (2) and (5). For, (5) is satisfied by the last statement, (2) implies (1), and (2) is satisfied by the definition of S(m). If one of the above requirements has to be violated then an extension of the above algorithm is needed; in particular, a modification of (ii) in Step B. An example of using the algorithm is shown in the appendix.

The sets S(m), defined at the beginning of this section, are constructed recursively as follows:

$$S(1) = \left\{R_k : c_{k1} = 1\right\}$$

.

.

.

$$S(m+1) = \left\{R_k : R_k \notin \bigcup_{p=1}^{m} S(p),\ c_{kj} = 1 \text{ for some } R_j \in S(m)\right\}$$

That is, to construct S(m+1) it is necessary to examine in the matrix C only the entries 1 in the columns which correspond to repeaters of S(m) and choose the ones that have not been identified yet.

III.3 Remarks

1. The shortest path label assignment does not, in general, correspond to physical distance. That is, the label assignment depends on the terrain as well as possible variations in transmission power and reception sensitivity of devices. Thus, it is a functional rather than a geophysical assignment.

2. In the practical case, it may be necessary to label redundent repeaters and then turn them off to be used as stand-by repeaters. Furthermore, the procedure for doing this should satisfy requirements (3) and (4) in the previous section. The study of such techniques can be done within the framework of reliability.

IV. REFERENCES

[1] Packet Radio - Systems Considerations; PRTN No. 5, Network Analysis Corporation, January 1, 1973.

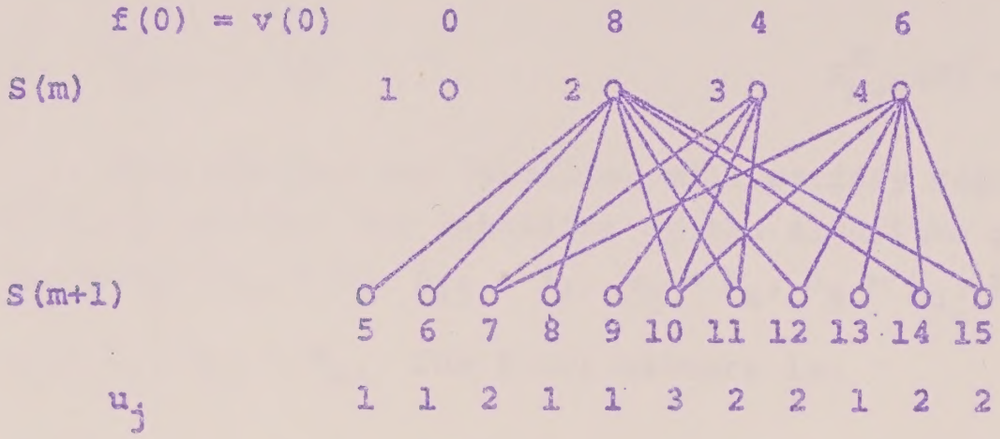
[2] Combinatorial Aspects of Flow in Packet Radio Nets; PRTN No. 6. Network Analysis Corporation, December 1972.

[3] Comparison of Hop-by-Hop and End-to-End Acknowledgment Schemes; PRTN No. 7, Network Analysis Corporation, January 12, 1973.

[4] L. G. Roberts: Extensions of Packet Communication Technology to a Hand Held Personal Terminal; Spring Joint Computer Conference, 1972.

APPENDIX

The figure below shows the set S(m+1) to be labelled by S(m). The connection line between $R_i \varepsilon S(m+1)$ and $R_j \varepsilon S(m)$ was drawn to demonstrate that $c_{ij} = 1$, also B = 3 bits.



$S_m^F(0) = \{R_1\}$ $\qquad$ $S_{m+1}^F(0) = \{\phi\}$

$S_m^L(0) = \{R_3, R_4\}$ $\qquad$ $S_{m+1}^L(0) = \{R_{10}, R_{11}, R_{12}, R_{14}, R_{15}, R_5, R_6, R_8\}$

$S_m^R(0) = \{R_2\}$ $\qquad$ $S_{m+1}^R(0) = \{R_7, R_9, R_{13}\}$

First label assigned is $h_{10} = R_3$, then,

$S_m^F(1) = \{R_1\}$ $\qquad$ $S_{m+1}^F(1) = \{R_{10}\}$

$S_m^L(1) = \{R_4, R_3\}$ $\qquad$ $S_{m+1}^L(1) = \{R_{11}, R_{12}, R_{14}, R_{15}, R_5, R_6, R_8\}$

$S_m^R(1) = \{R_2\}$ $\qquad$ $S_{m+1}^R(1) = \{R_7, R_9, R_{13}\}$

The second label is $h_{11} = R_3$, then,

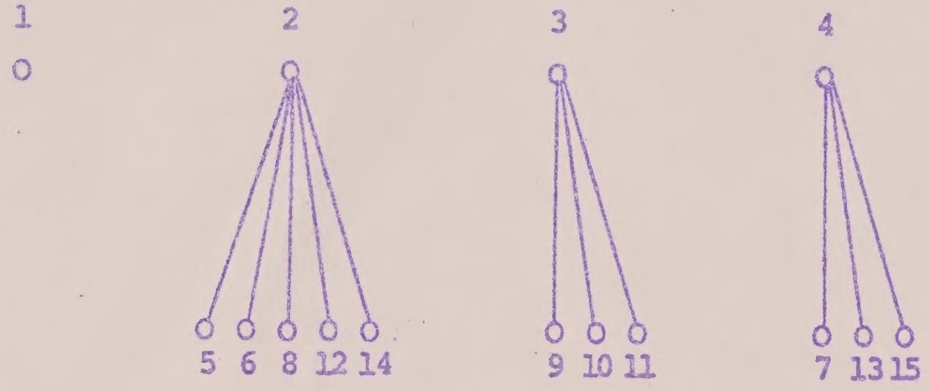
$S_m^F(2) = \{R_1\}$

$S_m^L(2) = \{R_2, R_4, R_3\}$

$S_m^R(2) = \{\phi\}$

$S_{m+1}^F(2) = \{R_{10}, R_{11}\}$

$S_{m+1}^L(2) = \{R_{11}, R_{12}, R_{14}, R_{15}, R_7, R_5, R_6, R_8, R_9, R_{13}\}$

$S_{m+1}^R(0) = \{\phi\}$

From now on, any lebelling will satisfy requirement (5) since $S_m^R(2)$ is empty. The labelling by the algorithm proceeds as follows: $h_{12} = R_2$, $h_{14} = R_2$, $h_{15} = R_4$, $h_7 = R_4$, $h_5 = R_2$, $h_6 = R_2$, $h_8 = R_2$, $h_9 = R_3$, $h_{13} = R_4$. The final network is: